

The Use of Research Evidence in Instructional Improvement

By Tom Corcoran

School district leaders want to make evidence-based decisions and they are making efforts to build evidence-based cultures in their central offices and schools. But, significant progress is being hampered by the inadequacy and confusion of the existing research, its availability to school and district-level staff, and a reliance by staff on decision-making patterns that focus on philosophy rather than effects.

Introduction

Those who seek to reform our public schools often argue that school performance would improve if only policy and practice were based on evidence. If decision-makers and practitioners paid more attention to research findings, the argument goes, they would make better decisions about improvement strategies and resource allocation, and we would see better results. The belief in this axiom is demonstrated by the increasing frequency with which reformers, educators, and policymakers find it necessary to legitimate their actions with claims that they are “research-based.” However, moving beyond rhetoric to actually put this principle into operation turns out to be difficult.

This issue of CPRE Policy Briefs looks at findings from the Consortium for Policy Research in Education’s study of how central office and school staff in three urban districts made decisions about instructional improvement strategies, and how much weight they gave to evidence. The three districts, in three different states, had enrollments ranging from 50,000 to over 200,000. Leaders in all three districts were addressing the problems common to most urban districts: students’

problems associated with living in poverty, low achievement, high mobility, and high dropout rates. Changes in district and school leadership, high teacher turnover, changes in funding, and new state policies compounded the difficulty of improving performance in all three sites. The districts were also working in environments characterized by decentralized decision-making, high-stakes accountability, and increasing competition among providers of comprehensive school reform designs and other “research-based” instructional improvement strategies.

During our study, we looked at three sets of strategic decisions that each district faced as it tried to improve student performance. The first set of decisions concerned what to do (i.e., which reform design or curriculum to adopt). The next set was deciding how to get it done; how to provide adequate support and coordination; and how to focus people’s attention on the desired changes, ensure effective implementation, reduce distractions, and buffer this important work from competing agendas. Finally, there were decisions about scaling up the reforms — the problem of replication. The decision-making process in each case was complicated and the use of evidence to support the ultimate decisions varied considerably. While these categories overlap, they are useful organizers of our findings.

Consortium for Policy Research in Education

University
of Pennsylvania

Harvard University

Stanford University

University
of Michigan

University of
Wisconsin-Madison

Decision One: What to Do and How to Begin?

Use of Evidence

In these highly decentralized systems, central office staff in all three districts faced the challenge of deciding what their role should be in the design selection process. All three staffs wanted to ensure that school staffs made decisions about improvement strategies based on evidence that they were effective and adopted reforms that were considered “best practice.” However, finding ways to help schools use evidence-based decision-making proved to be difficult in all three situations.

One district set up a screening process that examined the alignment between local and state standards and whole-school designs with externally developed curricula and then reviewed the evidence supporting the claims of the developers. Schools could then choose from designs that had been approved by this district-wide committee. This process worked reasonably well for deciding which comprehensive school reform designs to adopt. The district committee narrowed the range of choices to a handful of designs and ruled out others whose programs were judged to be poorly aligned with state assessments. But this critical review of evidence was not extended to other reform initiatives or to investments in professional development.

A second district created an office dedicated to the review of evidence and planned to prepare and disseminate summaries of the research evidence on various school designs and curriculum programs so school staff could make more informed choices. But a central office staff committee created to guide this work became bogged down in ideological debates over what constituted evidence and whether specific programs should be included or excluded from the information base. When the available evidence did not correspond to their preferences, advocates of particular strategies attacked the quality of the evidence. A frequent argument heard was that standardized test scores were not an adequate measure of a program’s effectiveness. There was confusion about what it meant to be research-based — was a program

The Consortium for Policy Research in Education (CPRE) is funded by the Institute of Education Sciences, United States Department of Education under Grant No. R308A960003. Opinions expressed in this Brief are those of the author and do not necessarily reflect the views of the Institute of Education Sciences; the United States Department of Education; CPRE; or its institutional members.

research-based if research findings were used to determine its components or was evidence of effects necessary? The committee struggled over “rules of evidence.” Was evidence of effects in one school compelling?

In the third district, a committee of faculty from local higher education institutions and district staff was formed to screen designs and programs that would be permitted to enter the district, but the committee ultimately perceived their screening role as too prescriptive. They were reluctant to exclude any design or program on the grounds of inadequate evidence of its effectiveness because how one interpreted the available evidence was a matter of personal philosophy and perspective. They did not believe that decision rules could be developed that would be accepted by all parties. The committee shifted its focus from reviewing designs and programs to discussing partnerships among institutions.

In the end, patterns of decision-making based on philosophical commitments, political necessities, and the attractiveness or popularity of ideas prevailed over efforts to attend to evidence in all three districts. The emergence of evidence-based decision-making was hampered by professional cultures in which the “good” and the “popular” were valued more than the effective, and in which the lack of external sources of professional legitimacy meant that personal beliefs about policy and practice usually prevailed over evidence. But the effort to shift the focus to evidence was also hampered by difficulties accessing and making sense of the research evidence. An inadequate knowledge base, the lack of research synthesis, difficulty accessing the most recent research findings, and the ideological orientations of the researchers and consultants who were contacted all made it difficult for local decision-makers to attend to the evidence.

As a consequence, efforts to develop new procedures for review of evidence and to provide guidance to school staffs about “best practice” failed in two of the districts and were only partially successful in the third. Nevertheless, these efforts were important. They demonstrated a genuine desire on the part of district leaders to link their decision-making to the knowledge base about teaching and learning. They also indicated awareness on the part of these leaders that the existing organizational structures could not make this link, and that new mechanisms and new relationships were needed in order to make a shift to evidence-based practice.

Local Designs

During this period, each of the three districts designed their own reform initiatives that were intended to be widely, if not universally, adopted by their schools. Decisions about these local initiatives fell outside the scope of the new structures created to review evidence about best practice. However, the districts recognized that they were engaged in development work and that their new initiatives would have to be piloted and evaluated. They attempted to draw upon current knowledge and effective programs being used in other settings in designing their initiatives, but their use of research evidence in these instances was uneven. The champions of specific reforms typically examined literature selectively and found theories and “evidence” to justify their approaches or recruited “experts” who were advocates of their preferred strategies. Distinctions among empirical research, theories, and simple advocacy were not always clear or understood and for the reasons cited earlier, research evidence had limited influence on these local design processes.

Yet, the efforts to apply research to the design of new initiatives also revealed the seriousness of the local commitment to evidence-based practice. To ratchet up performance in its lowest performing schools, one district developed a new elementary school design that drew on elements of several nationally available programs. Another district drew on experience in industry to redesign its school organization and develop an innovative team approach to teaching. The third district had earlier piloted a strate-

gy for breaking up its large comprehensive high schools into small learning communities and moved to implement the strategy district-wide.

All three of these local initiatives were evaluated by external groups. The districts made significant investments in these evaluations, and all three used the results to make revisions in their designs. In two cases, the evaluations affected decisions about continuing or expanding the investments in these programs. District staff seemed to be more insistent that there be evidence of effectiveness for locally designed programs than they were about externally developed ones. This difference in attitude toward locally and externally developed programs may be related to the fact that the latter often brought external funds, and carried the weight of their national reputations. It might also be due to the greater entanglement of local initiatives in the internal politics of the systems and their association with specific actors or offices within the systems.

Central Office Versus School Staff

Central office staff showed a greater commitment to the use of evidence than did school staff. District staff generally felt that decisions ought to be based on a solid rationale supported by research or at least consistent with it. However, they were often frustrated by the lack of research on key issues with which they were struggling, such as the effectiveness of middle schools, high school literacy programs, and coaching strategies. In addition, the existing research they did find was often contradictory and confusing and it was hard to assess the claims made by various reform developers. School staff talked about using research, but in fact relied most heavily on the opinions of their peers. They found research hard to access and even harder to interpret and struggled to sort out significant findings from other knowledge claims.

Decentralization of decision-making combined with weak district guidance appeared to be undermining the use of knowledge rather than promoting it. Selection of whole-school reforms by school staffs was often based more on personal testimonials, philosophical comfort, ease of use, lack of threat to current practice, and good marketing by

developers than it was evidence of effects. Teachers indicated that research was less useful for determining if a design was producing results than their sense of whether the design fit their own notions of good practice. Teachers preferred designs that drew on research about practices that they already felt were “good” than they were in choosing designs that were producing results.

Impact of Time on Decisions

The pressure from local and state accountability systems, from civic leaders, and from parents to do “something” to raise student scores expeditiously had an enormous impact on how the districts proceeded. District staff understood the magnitude of their responsibility and felt the pressure to act quickly and to produce results. Leaders in all three districts felt that they had limited time, maybe two to three years, to produce gains in student performance. They hoped that increased performance pressure and more professional development for teachers would lead to changes in instruction and curriculum that would improve test scores in the short run and buy time for school designs and local initiatives to spread and have some effect. They were forced to allocate scarce resources to “short-run” strategies in order to buy time for the implementation and rollout of the deeper reforms that they were advocating.

This made it hard to maintain focus. In two cases, this strategy failed; district-wide test scores remained relatively flat. Gains in the schools implementing the local reform initiatives were not able to overcome the public perception that little progress was being made. In the third case, modest district-wide achievement gains stalled after several years and district-led reforms were undermined by weak implementation and resistance in the schools. In all three cases, the superintendents who had led the design of the local reform initiatives left under pressure from their boards and local political leaders who were growing impatient. In two cases, they were pushed out in spite of demonstrated growth in student achievement in the sites implementing reforms and new leadership transitions brought new agendas that distracted attention away from the current reforms.

Decision Two: How to Get the Job Done?

Professional Development

Leaders in all three districts believed that high-quality, stable professional development for teachers was essential to the effective implementation of their reform, and they believed that professional development should be research-based. Each district created new structures for this purpose.

One district created a quasi-independent academy that offered short, one- or two-day courses for teachers that were not subject-based, but focused instead on generic teaching competencies. These courses were based on research on instruction, but there was little attention to how curriculum content influences teaching and learning. These courses were popular with teachers and were seen as particularly useful by new teachers. The academy also provided training in support of the district reform initiative, but not for the comprehensive school reform designs. The latter task was left to the reform developers. There was no follow-up support for the training offered by the academy and observations in the schools suggested that it had limited impact on practice. The independent status of the academy made it harder for the district to redirect professional development to support reform initiatives and to improve its impact on instruction.

A second district had a staff of trainers who annually created a broad menu of workshops and courses in cooperation with schools and local universities. Some of the offerings related to the district reforms, especially standards-based instruction in various disciplines, but others covered a wide variety of topics. While many of the workshops were described as research-based, there was no procedure for determining whether these claims were accurate. Professional development was largely determined by the preferences of individual teachers, and there was no follow-up to support changes in practice. The district attempted to create a new model in which one school campus would serve as a demonstration site for new strategies and would work with networks of schools to test and diffuse these practices. The burden of running a large campus overwhelmed the

vision of research and development and the strategy never took hold. The central office was unwilling or unable to focus its professional development investments in support of its reforms. The culture of responding to teacher interests proved resilient.

The third district developed a large network of full-time professional development specialists to run workshops and coach teachers in support of the district reforms. Although intended to be a resource to the schools and responsive to their needs, they were more often assigned to support district reform initiatives. Much of their time was spent providing informational workshops for teachers. When they were able to work directly with teachers, they seemed to have a positive impact on practice. They had little connection to the comprehensive school reform design teams working in the schools. This district belatedly recognized that this network of trainers could not provide teachers with the knowledge and skills they needed to develop new curricula based on the newly adopted standards. Two years after the standards were adopted, the district began to offer curriculum summer institutes for teachers. Enrollment in summer programs grew rapidly and they were perceived by teachers as highly effective. However, follow-up support was uneven and the relationship between the institutes and the professional development network was weak.

The norms that shaped professional development in these districts proved hard to change. All three districts created professional development infrastructures that proved hard to control and to focus. They relied heavily on “trainers” who could design and deliver professional development and were not successful in moving this work into the schools. They had difficulty in developing strategies that addressed curriculum content. All three left fundamental decisions about the focus of professional development to individual teachers and, to a lesser degree, to schools. The result in all three cases was a highly fragmented system of professional development that failed to establish powerful connections to classroom practice.

Use of Evidence

Research evidence had little to do with the professional development offerings in any of the districts. The professional development staff were not members of an evidence-based culture themselves and they were seldom asked by district leadership if participation in the programs led to changes in practice or increases in student performance. There was no monitoring of the effects of professional development activities in any of the districts and no information gathered from teachers about any impact on their teaching practice. Evaluation was limited to satisfaction surveys. Since professional development staff felt they would be judged on how well they were liked, they often offered hot topics of the day in hopes of drawing and pleasing more teachers. Also, while the professional development staff were skilled as trainers and in group management, they had little expertise in substantive areas and were usually not linked to the research community. Professional development was driven by self-defined needs and by interest, and not by evidence of effectiveness.

Effectiveness

There were several reasons why districts had trouble offering effective professional development. One was that the districts themselves were not focused and were supporting many initiatives at the same time and they expected professional development staff to support them all. Too often, professional development time was used to transfer central office information and not focused on substantive issues. The professional development cultures themselves were not evidence-based and were instead committed to meeting the wants, needs, and desires of teachers, or pursuing the interests of the trainers. This latter point is important. The reliance on professional trainers meant that there were jobs to be protected. The trainers had a vested interest in offering the kinds of professional development experiences that they knew how to design and deliver. This culture was entrenched and hard to change. And finally, most professional development decisions were being made at levels below the central office.

District- and school-sponsored professional development activities most often neglected the content knowledge of teachers. Programs tended to emphasize process and procedure over subject-matter knowledge. The staff responsible for this enormous task were linked to different reform networks and valued different approaches to improvement. The result was a potpourri of workshops and events rather than a coherent program of professional development.

Decision Three: How to Scale Up?

All three districts made attempts to take reforms to scale. Two of the districts relied heavily on incentives to persuade schools to adopt reforms, and the third used mandates. However, little thought was given to how teachers viewed the new reform programs, how to persuade any reluctant staff to adopt new practices, or how to help them succeed. Many central office staff believed that acceptance and expansion of reforms would happen by itself, spurred on by the pressure to improve performance that the high-stakes accountability environment would bring to bear. The pressure would lead to the adoption of “proven” practices and, over time, school staff would make better adoption decisions.

In fact, there is some evidence to support this belief. Literacy programs spread rapidly across two of the districts because two schools that adopted the programs early showed higher gains on reading tests than other schools. But, there is contrary evidence as well. In one district, a locally designed elementary program proved highly successful in a set of low-performing schools, but it was not adopted by other schools in spite of the central office’s effort to provide incentives for its adoption. Inquiry suggested that it was seen as a “remedial” strategy for low-income, at-risk students and its adoption would be viewed by teachers and parents as an admission of failure by school leaders. Schools often continued to participate in some designs even after several years with little or no gains. Also, district accountability systems placed a premium on obtaining quick effects and reforms require several years to achieve any significant change in classroom practice.

The common tendencies of the districts were to either offer opportunities to schools or to mandate reforms and provide the necessary training. Concerns about implementation barriers, important incentives or disincentives, and legitimate alternatives were not seriously discussed or considered. Using evidence in the decision to expand was not common. In fact, decisions to expand a program often ran ahead of the collection and review of evidence. Again, time pressure proved more powerful than the desire to base decisions on evidence.

There were several external forces that made scaling up difficult for districts — the noisy reform environment, changes in state policies, and turnover in leadership. Focusing investments, persisting to get full implementation across a district, and making investment and continuation decisions in the face of changes in central office leadership was almost impossible. The leadership turnover at the middle level was also significant and critical because so often it is the staff at this level that reviews the evidence and advocates for specific programs. District leaders wanted to use evidence as they made decisions about whether to scale up their reform. But this was in conflict with their need to appear knowledgeable and confident about their vision and strategy in the eyes of their board and the public. And when they did support evaluations and feedback, it came too late. They couldn’t wait for years to decide next steps and so they made decisions to replicate the reform lacking the solid evidence most really wanted. Scaling up happened more by default than by plan and was seldom based on evidence of a program’s success.

Conclusion

The fact that central office personnel in the three districts were seeking research findings represents an important step toward evidence-based decision-making. Leaders in all three districts said that they wanted staff to base their decisions on evidence whenever possible. They also wanted to support and expand programs that produced results. In all three districts, there were serious efforts to build evidence-based cultures in the central offices and to encourage schools to pay greater attention to research evidence. How-

ever, these efforts were seriously hampered by the limited research evidence easily available to decision-makers and by problems faced by staff trying to access and make sense of the research they could find. Conflicting research findings, small sample sizes, and a lack of attention to key issues facing districts all contributed to the confusion. Districts were also hampered by the inability of district and school staff to put aside old patterns of decision-making that focused on philosophy or the “goodness” of an option, rather than evidence of its effectiveness.

Nevertheless, some progress was made as central office staff gave more consideration to research evidence in their decision-making. There was much less progress, however, in shifting the mindsets of school staff or in persuading professional development staff to use evidence, not interest, as the primary criterion for development programs for teachers. The leaders in the three districts agreed that most investment decisions were not being made on the evidence and viewed this as a significant weakness, but they also felt that correcting it would require major cultural changes in their organizations.

Finally, while many current attempts to improve the use of evidence focus on the supply side of the equation (that is, on improving the quality and definitiveness of education research), this study shows the importance of the demand side as well. Quite simply, users must be educated to appreciate the value and contributions of research.

Next Steps

What can be done to help district leaders and school staffs move in the direction of evidence-based practice? Part of the answer lies in making changes in the priorities and approaches of the research community. More collaboration is needed between researchers and practitioners in setting research agendas. Research products need to be more accessible and rigorous. Likewise, districts must demand high-quality research that is based on well-designed and reviewed studies. Increased demand for useful and usable research products will influence what researchers do, but this will happen more quickly if funding agencies take on the task of focusing research on the core problems of practice. All parties must work to actively

reduce the noisiness of education research that now makes it so hard for local decision-makers to make sense of the message. Federal agencies need to support periodic meta-analyses in critical areas in order to make the knowledge base accessible to local decision-makers.

But changes in the research community must be paralleled with changes in the culture of school districts. The careerism and political climates that result in constant changes in leadership and in reform agendas are also major parts of the problem. The work of improving teaching and learning is serious work, and it needs to be treated as such. It is work that takes time and significant investments. It may be impossible to bring about more stable leadership, but it should be possible to stabilize state and local reform agendas. State policymakers, business leaders, civic leaders in districts, and board members need to be educated about the high cost of making constant changes in direction and of layering reforms over one another.

Furthermore, districts are well positioned to carry out large clinical studies themselves. Large districts could test new curriculum or instructional strategies. They could engage large numbers of teachers in clinical trials of specific instructional techniques. They should not be waiting for the academic research community to meet their needs, but should form partnerships with local research institutions and seek funding for research that will meet their needs. The key point is that they should lead this work, not be submissive partners in research directed by researchers whose interests may not match their own.

Finally, it is important not to oversell the importance of research findings or underestimate the value of local “clinical” knowledge. The evidence-based practice movement in medicine values both scientific research and clinical expertise. Good treatment rests on both sources of knowledge. This approach should be followed in education. Research findings can suggest policy options or components of a local response to a problem, but determining what will work best in a particular district requires the knowledge of its existing commitments, its history, its resources, and its political climate. Increased

attention to research findings and the clinical expertise of accomplished practitioners will lead to better decisions and to a strong knowledge base, but they are only part of what must be considered by policymakers committed to evidence-based policymaking.

About the Author

Tom Corcoran is co-director of the Consortium for Policy Research in Education (CPRE). Prior to joining CPRE, Corcoran served as the education policy advisor for former New Jersey Governor James Florio, director of school improvement at Research for Better Schools, and director of evaluation and later chief-of-staff of the New Jersey Department of Education. His research interests include the use of research to inform policy and practice in public education, the efficacy of different approaches to professional development, the effectiveness of whole-school reform, the impact of changes in work environments on the productivity of teachers and students, and the factors affecting the effectiveness of scaling-up strategies.

About CPRE

The Consortium for Policy Research in Education (CPRE) studies alternative approaches to education reform in order to determine how state and local policies can promote student learning. Currently, CPRE's work is focusing on accountability policies, efforts to build capacity at various levels within the education system, methods of allocating resources and compensating teachers, instructional improvement, finance, and student and teacher standards. The results of this research are shared with policymakers, educators, and other interested individuals and organizations in order to promote improvements in policy design and implementation.

CPRE unites five of the nation's leading research institutions to improve elementary and secondary education through research on policy, finance, school reform, and school governance. Members of CPRE are the University of Pennsylvania, Harvard University, Stanford University, the University of Michigan, and the University of Wisconsin-Madison.

CPRE Policy Briefs are published by CPRE. To learn more about CPRE research or publications, please call 215-573-0700 or access CPRE publications at www.cpre.org; www.wcer.wisc.edu/cpre/; or www.sii.soe.umich.edu.



Graduate School of Education
University of Pennsylvania
3440 Market Street, Suite 560
Philadelphia, PA 19104-3325

NON PROFIT
U.S. POSTAGE

PAID
PERMIT NO. 2563
PHILADELPHIA, PA
